

Internet Archive

From Wikipedia, the free encyclopedia

Coordinates: 37°46′56.3″N 122°28′17.65″W﻿ / ﻿

The **Internet Archive** is a non-profit digital library with the stated mission of "universal access to all knowledge."^{[2][3]} It provides permanent storage of and free public access to collections of digitized materials, including websites, music, moving images, and nearly three million public-domain books. As of October 2012, its collection topped 10 petabytes.^{[4][5]} In addition to its archiving function, the Archive is an activist organization, advocating for a free and open Internet.

The Internet Archive allows the public to upload and download digital material to its data cluster, but the bulk of its data is collected automatically by its web crawlers, which work to preserve as much of the public web as possible. Its web archive, The Wayback Machine, contains over 150+ billion web captures.^{[6][7]} The Archive also oversees one of the world's largest book digitization projects.

Founded by Brewster Kahle in 1996, the Archive is a 501(c)(3) non-profit operating in the United States. It has an annual budget of \$10 million, derived from a variety of sources: revenue from its Web crawling services, various partnerships, grants, donations, and the Kahle-Austin Foundation.^[8] Its headquarters are in San Francisco, California, where about 30 of its 200 employees work. Most of its staff work in its book-scanning centers. The Archive has data centers in San Francisco, Redwood City, and Richmond, all in California. Its collection is mirrored for stability and endurance at the Bibliotheca Alexandrina in Egypt.^[9]

The Archive is a member of the International Internet Preservation Consortium.^[10] This non-profit digital library was officially designated as a library by the State of California in 2007.^[11]

Contents

- 1 History
- 2 World Wide Web archiving
 - 2.1 Wayback Machine
 - 2.2 Archive-It
- 3 Books collections
 - 3.1 Text collection
 - 3.2 Number of texts for each language
 - 3.3 Number of texts for each decade
 - 3.4 Open Library
- 4 Media collections
 - 4.1 Moving image collection
 - 4.1.1 Machinima archive
 - 4.2 TV News Search & Borrow
 - 4.3 Audio collection
 - 4.4 Netlabels
 - 4.5 NASA Images
- 5 Other services and endeavors
 - 5.1 Physical media
 - 5.2 Software
- 6 Controversies and legal disputes
 - 6.1 *Omni* magazine
 - 6.2 Grateful Dead
 - 6.3 National security letter
 - 6.4 Uncensored hosting
 - 6.5 Opposition to Google Books settlement
 - 6.6 Opposition to SOPA and PIPA bills
- 7 Ceramic Archivists collection
- 8 List of digitizing sponsors for ebooks
- 9 See also
- 10 References
- 11 Further reading
- 12 External links

History


Kahle founded the Archive in 1996 at the same time that he began the for-profit web crawling company Alexa Internet. In 1996, The Internet Archive had begun to archive and preserve the World Wide Web. The archived content wasn't available until 2001, when it developed the Wayback Machine. In late 1999, the Archive expanded its collections beyond the Web archive, beginning with the Prelinger Archives. Now the Internet Archive includes texts, audio, moving images, and software. It hosts a number of other projects: the NASA Images Archive, the contract crawling service Archive-It, and the

Internet Archive	
<div><div>INTERNET</div><div>ARCHIVE</div><div></div></div>	
Type	501(c)(3) non-profit
Founded	1996
Headquarters	Richmond District, San Francisco, California, U.S.
Chairman	Brewster Kahle
Services	Archive-It, Open Library, Wayback Machine (since 2001), Netlabels, NASA Images, Prelinger Archives
Employees	200
Slogan(s)	Universal access to all knowledge
Website	Archive.org (https://archive.org)
Alexa rank	▼ 159 (April 2014) ^[1]
Type of site	Digital library
Available in	English
Launched	2001

Headquarters



Since 2009, headquarters have been at 300 Funston Avenue in San Francisco, a former Christian Science Church.



Mirror of the Internet Archive in the Bibliotheca Alexandrina

wiki-editable library catalog and book information site Open Library. Recently, the Archive has begun working to provide specialized services relating to the information access needs of the print-disabled; publicly accessible books were made available in a protected Digital Accessible Information System (DAISY) format.^[12]

According to its website:

Most societies place importance on preserving artifacts of their culture and heritage. Without such artifacts, civilization has no memory and no mechanism to learn from its successes and failures. Our culture now produces more and more artifacts in digital form. The Archive's mission is to help preserve those artifacts and create an Internet library for researchers, historians, and scholars.

In August 2012, the Archive announced^[13] that it has added BitTorrent to its file download options for over 1.3 million existing files, and all newly uploaded files.^{[14][15]} This method is the fastest means of downloading media from the Archive, as files are served from two Archive data centers, in addition to other torrent clients which have downloaded and continue to serve the files.^{[14][16]}

On November 6, 2013, the Internet Archive's headquarters in San Francisco's Richmond District caught fire, destroying equipment and damaging some nearby apartments.^[17] According to the Archive, it lost:^[18]

- a side-building housing one of 30 of its scanning centers
- cameras, lights, and scanning equipment worth hundreds of thousands of dollars
- "maybe 20 boxes of books and film, some irreplaceable, most already digitized, and some replaceable."

The non-profit Archive sought donations to cover the estimated \$600,000 in damages.^[19]

World Wide Web archiving

Wayback Machine

Main article: Wayback Machine

The Internet Archive has capitalized on the popular use of the term "WABAC Machine" from a segment of the old Rocky and Bullwinkle cartoon, and uses the name "Wayback Machine" for its service that allows archives of the World Wide Web to be searched and accessed.^[20] This service allows users to view archived web pages. The Wayback Machine was created as a joint effort between Alexa Internet and the Internet Archive when a three-dimensional index was built to allow for the browsing of archived web content.^[21] Millions of websites and their associated data (images, source code, documents, etc.) are saved in a gigantic database. The service can be used to see what previous versions of websites used to look like, to grab original source code from websites that may no longer be directly available, or to visit websites that no longer even exist. The Internet Archive Terms of Use specify that users of the Wayback Machine are not to download data from the collection. Not all websites are available because many website owners choose to exclude their sites. As with all sites based on data from web crawlers, the Internet Archive misses large areas of the web for a variety of other reasons. International biases have also been found in its coverage, although this does not seem to be the result of a deliberate policy.^[22]

The use of the term "Wayback Machine" in the context of the Internet Archive has become so common that "Wayback Machine" and "Internet Archive" are almost synonymous. This usage occurs in popular culture, e.g., in the television show *Law and Order: Criminal Intent* ("Legacy", first run Aug 3, 2008), an extra playing a computer tech uses the "Wayback Machine" to find an archive of a student's Facebook style website. Snapshots usually take at least 6–18 months to be added.

The Save Page Now archiving feature was made available in October 2013,^[23] accessible on the lower right of the Wayback Machine's main page.^[24] Once a target URL is entered and saved, if the target website permits access via robots.txt, the web page will become part of the Wayback Machine.^[23]

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013
Number of all archived pages (billion)	40 ^[25]	85 ^[26]	85 ^[27]	85 ^[28]	150 ^[29]	150 ^[30]	150 ^[31]	150 ^[32]	373 ^[33]

Archive-It

Created in early 2006, Archive-It^[34] is a web archiving subscription service that allows institutions and individuals to build and preserve collections of digital content and create digital archives. Archive-It allows the user the option to customize their capture or exclusion of web content they want to preserve for cultural heritage reasons. Through a web application, Archive-It partners can harvest, catalog, manage, browse, search and view their archived collections. In terms of accessibility, the archived websites are full text searchable within seven days of capture.^[35] Content collected through Archive-It is captured and stored as a WARC file. A primary and back-up copy is stored at the Internet Archive data centers. A copy of the WARC file can be given to subscribing partner institutions for geo-redundant preservation and storage purposes to their best practice standards.^[36] The data captured through Archive-It is periodically indexed into the Internet Archive's general archive.



From 1996 to 2009, headquarters were in the Presidio of San Francisco, a former U.S. military base.



Wayback Machine logo, used since October 24, 2001



A purchase of additional storage at the Internet Archive



Brewster Kahle of the Internet Archive talks about archiving operations

As of March 2014, Archive-It had over 275 partner institutions in 46 U.S. States and 16 countries that have captured over 7.4 billion URLs for over 2444 public collections. Archive-It partners are universities and college libraries, state archives, federal institutions, museums, law libraries and cultural organizations, including the Electronic Literature Organization, North Carolina State Archives and Library, Stanford University, Columbia University, American University in Cairo, Georgetown Law Library and many others.

Books collections

Text collection

The *Text* (<http://www.archive.org/details/texts>) collection at the **Internet Archive** ^[more]

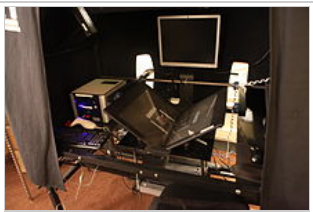
The Internet Archive Text Archive collection includes digitized books and special collections from various libraries and cultural heritage institutions from around the world.

The Internet Archive operates 33 scanning centers in five countries, digitizing about 1,000 books a day for a total of over 2 million books,^[37] financially supported by libraries and foundations.^[38] As of July 2013, the collection included 4.4 million books with over 15 million downloads per month.^[37] As of November 2008, when there were about 1 million texts, the entire collection was over 0.5 petabytes, which includes raw camera images, cropped and skewed images, PDFs, and raw OCR data.^[39]

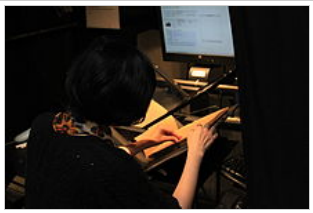
Between about 2006 and 2008 Microsoft Corporation had a special relationship with Internet Archive texts through its Live Search Books project, scanning over 300,000 books which were contributed to the collection, as well as financial support and scanning equipment. On May 23, 2008, Microsoft announced it would be ending the Live Book Search project and no longer scanning books.^[40] Microsoft made its scanned books available without contractual restriction and donated its scanning equipment to its former partners.^[40]

Around October 2007 Archive users began uploading public domain books from Google Book Search.^[41] As of November 2013 there were over 900,000 Google-digitized books in the Archive's collection;^[42] the books are identical to the copies found on Google, except without the Google watermarks, and are available for unrestricted use and download.^[43]

Besides books, the Archive offers free and anonymous public access to more than a million court opinions, legal briefs, or exhibits uploaded from the United States Federal Courts' PACER electronic document system via the RECAP web browser plugin. All of these documents are in the public domain, but had been kept from the public behind a federal court paywall. The Archive has copies of less than 1% of PACER's documents.



Internet Archive "Scribe" book scanning workstation



An Internet Archive in-house scan ongoing

Number of texts for each language

Number of all texts (March 30, 2014)	6,024,064 ^[44]
---	---------------------------

Language	English	French	German	Spanish	Russian	Arabic	Urdu	Portuguese	Chinese	Japanese	Dutch
Number of texts (January 21, 2014)	4,284,928 ^[45]	293,393 ^[46]	222,689 ^[47]	114,906 ^[48]	17,487 ^[49]	30,603 ^[50]	7,031 ^[51]	14,643 ^[52]	81,311 ^[53]	7,410 ^[54]	19,076 ^[55]

Number of texts for each decade

Decade	1800-1809	1810-1819	1820-1829	1830-1839	1840-1849	1850-1859	1860-1869	1870-1879	1880-1889	1890-1899
Number of texts (January 24, 2014)	31,024 ^[56]	41,619 ^[57]	62,856 ^[58]	82,676 ^[59]	99,891 ^[60]	140,260 ^[61]	163,908 ^[62]	182,258 ^[63]	248,371 ^[64]	320,526 ^[65]

Decade	1900-1909	1910-1919	1920-1929	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1989	1990-1999	2000-2009	2010s (decade)
Number of texts (January 24, 2014)	447,346 ^[66]	408,317 ^[67]	156,345 ^[68]	49,279 ^[69]	52,620 ^[70]	55,192 ^[71]	85,777 ^[72]	88,287 ^[73]	94,251 ^[74]	170,330 ^[75]	375,553 ^[76]	400,401 ^[77]

Open Library

Main article: Open Library

The Open Library is another project of the Internet Archive. The site seeks to include a web database for every book ever published: it holds 23 million catalog records of books. It also seeks to be a web-accessible public library: it contains the full texts of about 1,600,000 public domain books (out of the over five million from the main texts collection), which are fully readable, downloadable^{[78][79]} and full-text searchable;^[80] it offers access to an e-book lending program for over 250,000 recent books not in the public domain, in partnership with over 1,000 library partners from 6 countries^{[37][81]} (after getting a "library card", that is a free registration on the website).

Open Library is a free/open source software project, with its source code freely available on the Open Library site.

Media collections

In addition to web archives, the Internet Archive maintains extensive collections of digital media that are attested by the uploader to be in the public domain in the United States or licensed under a license that allows redistribution, such as Creative Commons licenses. Media are organized into collections by media type (moving images, audio, text, *etc.*), and into sub-collections by various criteria. Each of the main collections includes an "Community" sub-collection (formerly named "Open Source") where general contributions by the public are stored.

Moving image collection

The *Moving Image Archive* (<http://www.archive.org/details/movies>) collection at the **Internet Archive** ^[more]

The Internet Archive holds a collection of approximately 3,863 feature films.^[82] Additionally, the Internet Archive's Moving Image collection includes: newsreels, classic cartoons, pro- and anti-war propaganda, The Video Cellar Collection, Skip Elsheimer's "A.V. Geeks" collection, and ephemeral material from Prelinger Archives, such as advertising, educational, and industrial films and amateur and home movie collections.

Subcategories of this collection include:

- IA's *Brick Films* collection, which contains stop-motion animation filmed with Lego bricks, some of which are "remakes" of feature films.
- IA's *Election 2004* collection, a non-partisan public resource for sharing video materials related to the 2004 United States Presidential Election.
- IA's *FedFlix* collection, Joint Venture NTIS-1832 between the National Technical Information Service and Public.Resource.Org that features "the best movies of the United States Government, from training films to history, from our national parks to the U.S. Fire Academy and the Postal Inspectors"^[83]
- IA's *Independent News* collection, which includes sub-collections such as the Internet Archive's World At War competition from 2001, in which contestants created short films demonstrating "why access to history matters". Among their most-downloaded video files are eyewitness recordings of the devastating 2004 Indian Ocean earthquake.
- IA's *September 11th Television Archive*, which contains archival footage from the world's major television networks of the terrorist attacks of September 11, 2001 as they unfolded on live television.^[84]

Some of the films available on the Internet Archive are:^[85]

- *The 39 Steps* (1935)
- *Battleship Potemkin* (1925)
- *The Birth of a Nation* (1915)
- *Broken Blossoms* (1919)
- *The Century of the Self* (2002 documentary television series)
- *Charade* (1963)
- *Columbia Revolt* (1968)
- *D.O.A.* (1950)
- *Danger Lights* (1930)
- *Das Cabinet des Dr. Caligari* (1920)
- *Dating Do's and Don'ts* (1949)
- *Detour* (1945)
- *Duck and Cover* (1951)
- *Escape from Sobibor* (1987)
- *Fire Over England* (1937)
- *The General* (1927)
- *Greed* (1924)
- *Hemp for Victory* (1942)
- *Intolerance* (1916)
- *The Kid* (1921)
- *Le voyage dans la Lune* (1902)
- *Lying Lips* (1939)
- *M* (1931)
- *The Man Who Knew Too Much* (1934)
- *Manos: The Hands of Fate* (1966)
- *Manufacturing Consent: Noam Chomsky and the Media* (1992)
- *Night of the Living Dead* (1968)
- *Nosferatu* (1922; not public domain outside of the United States)
- *Phantom of the Opera* (1925)
- *Plan 9 from Outer Space* (1959)
- *Princess Iron Fan* (1941)
- *Reefer Madness* (1936)
- *Sex Madness* (1938)
- *She Done Him Wrong* (1933)
- *Triumph of the Will* (1935)
- *Why We Fight* (1942-1945; all seven documentary film episodes)

See also: *Wikipedia list of films freely available on the Internet Archive*



Media reader



Microfilms at the Internet Archive



Videocassettes at the Internet Archive

Machinima archive

One of the sub collections of the Internet Archive's Video Archive is the Machinima Archive. This small section hosts many Machinima videos (see Machinima: Virtual Filmmaking). Machinima is a digital artform in which computer games, game engines or software engine are used in a sandbox mode like mode to create motion pictures, recreate plays or even publish presentations/keynotes. The archive collects a range of Machinima films from internet publishers such as Rooster Teeth and Machinima.com as well as independent producers. The sub collection is a collaborative effort between the Internet Archive, the How They Got Game research project at Stanford University, the Academy of Machinima Arts and Sciences and Machinima.com.^[86]

TV News Search & Borrow

In September 2012, the Internet Archive launched the TV News Search & Borrow^[87] service for searching U.S. national news programs. The service is built on closed captioning transcripts and allows user to search and stream 30-second video clips. Upon launch, the service contained "350,000 news programs collected over 3 years from national U.S. networks and stations in San Francisco and Washington D.C."^[88] According to Kahle, the service was inspired by the Vanderbilt Television News Archive, a similar library of televised network news programs.^[89] In contrast to Vanderbilt, which limits access to streaming video to individuals associated with subscribing colleges and universities, the TV News Search & Borrow allows open access to its streaming video clips.



TV tuners at the Internet Archive

Audio collection

Main article: Live Music Archive

The *Audio Archive* (<http://www.archive.org/details/audio>) collection at the **Internet Archive** ^[*more*]

The Audio Archive includes music, audio books, news broadcasts, old time radio shows and a wide variety of other audio files. There are over 200,000 free digital recordings in the collection. The subcollections include audio books and poetry, podcasts, non-English audio and many others.^[90]

The Live Music Archive sub-collection includes over 100,000 concert recordings from independent artists, as well as more established artists and musical ensembles with permissive rules about recording their concerts such as the Grateful Dead, and more recently, The Smashing Pumpkins. Also, Jordan Zevon has allowed Internet Archive to host a definitive collection of his father Warren Zevon concert recordings. The catalog ranges from 1976-2001 and contains 1,137 free songs.^[91]

Netlabels

The Archive has a collection of freely distributable music that is streamed and available for download via its *Netlabels* service. The music in this collection generally have Creative Commons-license catalogs of virtual record labels.^{[92][93]}

NASA Images

The NASA Images archive was created through a Space Act Agreement between the Internet Archive and NASA to bring public access to NASA's image, video, and audio collections in a single, searchable resource. The IA NASA Images team worked closely with all of the NASA centers to keep adding to the ever-growing collection.^[94] The nasaimages.org site launched in July 2008 and had more than 100,000 items online at the end of its hosting in 2012.

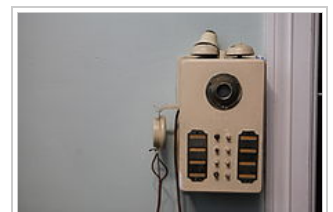
Other services and endeavors

Physical media

Voicing a strong reaction to the idea of books simply being thrown away, and inspired by the Svalbard Global Seed Vault, Kahle now envisions collecting one copy of every book ever published. "We're not going to get there, but that's our goal," he said. Alongside the books, Kahle plans to store the Internet Archive's old servers, which were replaced in 2010.^[95]

Software

The Internet Archive is "the largest collection of historical software online in the world", spanning 50 years of computer history in terabytes of computer magazines and journals, books, shareware discs, FTP websites, video games, etc.^[96] In 2013 the Internet Archive began to provide abandonware video games browser-playable via MESS, for instance the Atari 2600 game E.T. the Extra-Terrestrial.^[97]



Example of other "archived" item

Controversies and legal disputes

See also: Wayback Machine#Use in legal evidence

Omni magazine

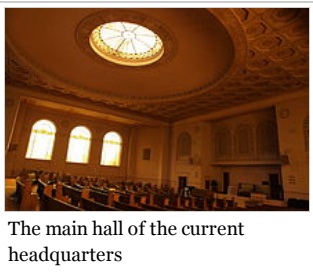
In a story at his Web site headed "What the heck is going on at Internet Archive?", author Steven Saylor noted, "Sometime in 2012, the entire run of Omni magazine was uploaded (and made available for download) at Internet Archive.... Since those old issues must contain hundreds of works still under copyright by numerous contributors, how is this legal?"^[98] At least one contributor to the magazine, author Steve Perry, has publicly

complained that he never gave permission for his work to be uploaded ("they didn't say a word in my direction"),^[99] and it has been noted that all issues containing the work of Harlan Ellison have apparently been taken down.^[100] Glenn Fleishman, investigating the question "Who Owns Omni?", writes that "Almost all of the authors, photographers, and artists whose work appeared in the magazine had signed contracts that granted only short-term rights....[No one] could simply reprint or post the content from older issues."^[101]

Grateful Dead

In November 2005, free downloads of Grateful Dead concerts were removed from the site. John Perry Barlow identified Bob Weir, Mickey Hart, and Bill Kreutzmann as the instigators of the change, according to a *New York Times* article.^[102] Phil Lesh commented on the change in a November 30, 2005, posting to his personal website:

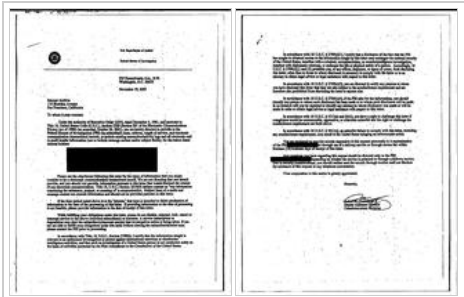
It was brought to my attention that all of the Grateful Dead shows were taken down from Archive.org right before Thanksgiving. I was not part of this decision making process and was not notified that the shows were to be pulled. I do feel that the music is the Grateful Dead's legacy and I hope that one way or another all of it is available for those who want it.^[103]



A November 30 forum post from Brewster Kahle summarized what appeared to be the compromise reached among the band members. Audience recordings could be downloaded or streamed, but soundboard recordings were to be available for streaming only. Concerts have since been re-added.^[104]

National security letter

On May 8, 2008, it was revealed that the Internet Archive successfully challenged an FBI National Security Letter asking for logs on an undisclosed user.^{[105][106]}



A National security letter issued to the Internet Archive demanding information about a user

Uncensored hosting

On August 17, 2011, Middle East Media Research Institute published "Al-Qaeda, Jihadis Infest the San Francisco, California-Based 'Internet Archive' Library"^[107] which detailed how members can post anonymously and enjoy free uncensored hosting.

Opposition to Google Books settlement

The Internet Archive is a member of the Open Book Alliance, which has been among the most outspoken critics of the Google Book Settlement. The Archive advocates an alternative digital library project.^[108]

Opposition to SOPA and PIPA bills

The Internet Archive blacked out its website for twelve hours on January 18, 2012 in protest of the Stop Online Piracy Act and the PROTECT IP Act bills, two pieces of pending legislation in the United States Congress that they claim will "negatively affect the ecosystem of web publishing that led to the emergence of the Internet Archive." This occurred in conjunction with the English Wikipedia blackout, as well as numerous other protests across the Internet.^[109]

Ceramic Archivists collection

The Great Room of the Internet Archive features a collection of nearly 100 ceramic figures representing employees of the Internet Archive. This collection was commissioned by Brewster Kahle and is ongoing. The sculptor is Nuala Creed (<http://www.internetarchivists.org>).



Ceramic figures of Internet Archive employees

List of digitizing sponsors for ebooks

This list is incomplete; you can help by expanding it
([//en.wikipedia.org/w/index.php?title=Internet_Archive&action=edit](http://en.wikipedia.org/w/index.php?title=Internet_Archive&action=edit)).

This is a list of some digitizing sponsors for ebooks in the Internet Archive.

Examples from the Wayback Machine's archives:

- Apple Computer (https://web.archive.org/web/*/www.apple.com)
- Amazon (https://web.archive.org/web/*/www.amazon.com)
- Microsoft (https://web.archive.org/web/*/www.microsoft.com)
- BBC News (https://web.archive.org/web/*/news.bbc.co.uk)
- Google (https://web.archive.org/web/*/google.com)
- Open Directory (https://web.archive.org/web/*/dmoz.org)

Sponsor	Collection	Number of texts (March 1, 2014)
Yahoo!	[1] (https://archive.org/details/yahoo_books)	1,076 ^[110]
Microsoft	[2] (https://archive.org/details/msn_books)	412,094 ^[111]
Google	[3] (https://archive.org/details/googlebooks)	907,760 ^[112]
Lyrasis Members and Sloan Foundation	[4] (https://archive.org/details/lyrasis)	85,845 ^[113]
University of Toronto	[5] (https://archive.org/details/university_of_toronto)	139,446 ^[114]
National Library of Scotland	[6] (https://archive.org/details/nationallibraryofscotland)	4,847 ^[115]
Natural History Museum Library, London	[7] (https://archive.org/details/nhml_london)	5,417 ^[116]
University of Alberta Libraries	[8] (https://archive.org/details/university_of_alberta_libraries)	76,472 ^[117]
Research Library, Getty Research Institute	[9] (https://archive.org/details/getty)	8,409 ^[118]
Boston Library Consortium	[10] (https://archive.org/details/blc)	37,482 ^[119]
The Library of Congress	[11] (https://archive.org/details/library_of_congress)	73,693 ^[120]
Allen County Public Library	[12] (https://archive.org/details/allen_county)	21,986 ^[121]
Internet Archive	[13] (https://archive.org/details/internetarchivebooks)	119,776 ^[122]
Harvard University	[14] (https://archive.org/details/Harvard_University)	7,805 ^[123]
China-America Digital Academic Library (CADAL)	[15] (https://archive.org/details/cadal)	78,371 ^[124]
University of Illinois Urbana-Champaign	[16] (https://archive.org/details/university_of_illinois_urbana-champaign)	53,076 ^[125]
University of North Carolina at Chapel Hill	[17] (https://archive.org/details/unclibraries)	18,639 ^[126]
Biodiversity Heritage Library	[18] (https://archive.org/details/biodiversity)	10,001 ^[127]

See also

- Lists of Internet Archive's collections

Similar projects

- Internet Memory Foundation
- Library of Congress Digital Library project
- National Digital Information Infrastructure and Preservation Program
- Project Gutenberg
- UK Government Web Archive at The National Archives (United Kingdom)
- UK Web Archiving Consortium
- WebCite

Other

- Archive Team
- Digital Dark Age
- Digital preservation
- Heritrix
- Link rot
- Memory hole
- PetaBox
- Wayback Machine
- Web archiving
- Web crawler

References

- ↑ "Archive.org Site Info" (<http://www.alexa.com/siteinfo/archive.org>). Alexa Internet. Retrieved 2014-04-01.
- ↑ "Internet Archive Frequently Asked Questions" (<http://www.archive.org/about/faqs.php#296>). Internet Archive. Retrieved April 13, 2013.
- ↑ "Internet Archive: Universal Access to all Knowledge" (<http://www.archive.org/details/SDForumBK>). Internet Archive. Retrieved April 13, 2013.
- ↑ "10,000,000,000,000 bytes archived!" (<http://blog.archive.org/2012/10/26/1000000000000000-bytes-archived/>). *Internet Archive Blogs*. October 26, 2012. "On Thursday, 25 October, hundreds of Internet Archive supporters, volunteers, and staff celebrated addition of the 10,000,000,000,000,000th byte to the Archive's massive collections."
- ↑ Brown, A. (2006). *Archiving websites: A practical guide for information management professionals*. London: Facet Publishing. p. 9.
- ↑ "Internet Archive: Projects" (<http://archive.org/projects/>). Internet Archive. Retrieved March 6, 2013.
- ↑ Grotke, A. (December 2011). "Web Archiving at the Library of Congress" (<http://www.infoday.com/cilmag/dec11/Grotke.shtml>). *Computers In Libraries*, v.31 n.10, p. 15-19. Information Today.
- ↑ Womack, David (Spring 2003). "Who Owns History?" (<http://www.cabinetmagazine.org/issues/10/womack.php>). *Cabinet Magazine* (10).
- ↑ "Donation to the new Library of Alexandria in Egypt" (<https://archive.org/about/bibalex.php>); Alexandria, Egypt; April 20, 2002. Bibliotheca Alexandrina (https://archive.org/about/bibalex_p_r.php). Internet Archive.
- ↑ "Members" (<https://web.archive.org/web/20100613021711/http://netpreserve.org/about/memberList.php>) at the Wayback Machine (archived June 13, 2010) International Internet Preservation Consortium. Netpreserve.org
- ↑ "Internet Archive officially a library" (<https://archive.org/iathreads/post-view.php?id=121377>), May 2, 2007. Internet Archive
- ↑ "Daisy Books for the Print Disabled" (<https://archive.org/details/printdisabled>), February 25, 2013. Internet Archive.
- ↑ Kahle, Brewster (August 7, 2012). "Over 1,000,000 Torrents of Downloadable Books, Music, and Movies" (<http://blog.archive.org/2012/08/07/over-1000000-torrents-of-downloadable-books-music-and-movies/>). *Internet Archive Blogs*.
- ↑ ***a b*** "Internet Archive Starts Seeding 1,398,875 Torrents" (<https://torrentfreak.com/internet-archive-starts-seeding-1398635-torrents-120807/>). TorrentFreak. August 7, 2012.
- ↑ "Hot List for bt1.us.archive.org (Updated August 7 2012, 7:31 pm PDT)" (<http://bt1.archive.org/hotlist.php>). *US Cluster*. Internet Archive.
- ↑ "Welcome to Archive torrents" (<https://archive.org/details/bittorrent>). Internet Archive.

17. ^ Alexander, Kurtis (November 6, 2013). "Internet Archive's S.F. office damaged in fire" (<http://www.sfgate.com/bayarea/article/Internet-Archive-s-S-F-office-damaged-in-fire-4960703.php>). San Francisco Chronicle.
18. ^ "Fire Update: Lost Many Cameras, 20 Boxes. No One Hurt" (<http://blog.archive.org/2013/11/06/scanning-center-fire-please-help-rebuild/>). *Internet Archive Blogs*. November 6, 2013.
19. ^ Shu, Catherine (November 6, 2013). "Internet Archive Seeking Donations To Rebuild Its Fire-Damaged Scanning Center" (<http://techcrunch.com/2013/11/06/internet-archive-seeking-donations-to-rebuild-its-fire-damaged-scanning-center/>). TechCrunch.
20. ^ Green, Heather (February 28, 2002). "A Library as Big as the World" (http://www.businessweek.com/technology/content/feb2002/tc20020228_1080.htm). Business Week Online.
21. ^ "Internet Archive. (2012). Frequently Asked Questions" (<http://www.archive.org/about/faqs.php>). Internet Archive. Retrieved April 13, 2013.
22. ^ Thelwall, Mike; Vaughan, Liwen (Spring 2004). "A fair history of the Web? Examining country balance in the Internet Archive" (http://www.scit.wlv.ac.uk/~cm1993/papers/fair_history_preprint.pdf). *Library & Information Science Research* **26** (2): 162–176. doi:10.1016/j.lisr.2003.12.009 (<http://dx.doi.org/10.1016%2Fj.lisr.2003.12.009>).
23. ^ **a b** Rossi, Alexis (October 25, 2013). "Fixing Broken Links on the Internet" (<http://blog.archive.org/2013/10/25/fixing-broken-links/>). Internet Archive. Retrieved December 29, 2013.
24. ^ "Wayback Machine main page" (<http://archive.org/web/>). Internet Archive. Retrieved December 30, 2013.
25. ^ "Internet Archive" (<https://web.archive.org/web/20051231080301/http://www.archive.org/index.php>). Internet Archive. Retrieved March 2, 2014.
26. ^ "Internet Archive" (<https://web.archive.org/web/20061228011056/http://www.archive.org/index.php>). Internet Archive. Retrieved March 2, 2014.
27. ^ "Internet Archive" (<https://web.archive.org/web/20071228170611/http://www.archive.org/index.php>). Internet Archive. Retrieved March 2, 2014.
28. ^ "Internet Archive" (<https://web.archive.org/web/20081224073445/http://www.archive.org/index.php>). Internet Archive. Retrieved March 2, 2014.
29. ^ "Internet Archive" (<https://web.archive.org/web/20091220201119/http://www.archive.org/index.php>). Internet Archive. Retrieved March 2, 2014.
30. ^ "Internet Archive" (<https://web.archive.org/web/20101230100945/http://www.archive.org/index.php>). Internet Archive. Retrieved March 2, 2014.
31. ^ "Internet Archive" (<https://web.archive.org/web/20110830094047/http://www.archive.org/index.php>). Internet Archive. Retrieved March 2, 2014.
32. ^ "Internet Archive" (<https://web.archive.org/web/20121231094801/http://archive.org/index.php>). Internet Archive. Retrieved March 2, 2014.
33. ^ "Internet Archive" (<https://web.archive.org/web/20131231032705/https://archive.org/>). Internet Archive. Retrieved March 2, 2014.
34. ^ "archive-it.org" (<http://www.archive-it.org/>). archive-it.org. Retrieved April 13, 2013.
35. ^ "What is the Difference between the General Archive (sometimes called the Wayback Machine) and Archive-It?" (<https://web.archive.org/jira.com/wiki/display/ARIH/Archive-It+How-to+FAQ#Archive-ItHow-toFAQ-differencebetween>). *Archive-It How to FAQ*. Archive-It. -- via Jira.com.
36. ^ "About Archive-It" (<http://www.archive-it.org/learn-more>). Archive-It. Retrieved March 3, 2014.
37. ^ **a b c** Hoffelder, Nate (July 9, 2013). "Internet Archive Now Hosts 4.4 Million eBooks, Sees 15 Million eBooks Downloaded Each Month" (<http://www.the-digital-reader.com/2013/07/09/internet-archive-now-hosts-4-4-million-ebooks-sees-15-million-ebooks-downloaded-each-month/>). The Digital Reader.
38. ^ Kahle, Brewster (May 23, 2008). "Books Scanning to be Publicly Funded" (<https://archive.org/iathreads/post-view.php?id=194217>). Internet Archive Forums.
39. ^ "Bulk Access to OCR for 1 Million Books" (<http://blog.openlibrary.org/2008/11/24/bulk-access-to-ocr-for-1-million-books/>). *Open Library Blog*. November 24, 2008.
40. ^ **a b** "Book search winding down" (<http://web.archive.org/web/20080820220749/http://blogs.msdn.com/livesearch/archive/2008/05/23/book-search-winding-down.aspx>). *MSDN Live Search Blog*. May 23, 2008. Archived from the original (<http://blogs.msdn.com/livesearch/archive/2008/05/23/book-search-winding-down.aspx>) on 2008-08-20.
41. ^ "Google Books at Internet Archive" (<https://archive.org/details/googlebooks>). Internet Archive.
42. ^ "List of Google scans" (<https://archive.org/search.php?query=sponsor%3A%28Google%29>) (search). Internet Archive.
43. ^ Books imported from Google have a metadata tag of scanner:google for searching purposes. The archive provides a link to Google for PDF copies, but also maintains a local PDF copy, which is viewable under the "All Files: HTTP" link.
44. ^ "Internet Archive Search : mediatype:texts" (<https://archive.org/search.php?query=mediatype%3Atexts>). Internet Archive. Retrieved March 30, 2014.
45. ^ "Internet Archive Search : (language:eng OR language:"English")" ([https://archive.org/search.php?query=\(language:eng%20OR%20language:%22English%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:eng%20OR%20language:%22English%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
46. ^ "Internet Archive Search : (language:fre OR language:"French")" ([https://archive.org/search.php?query=\(language:fre%20OR%20language:%22French%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:fre%20OR%20language:%22French%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
47. ^ "Internet Archive Search : (language:ger OR language:"German")" ([https://archive.org/search.php?query=\(language:ger%20OR%20language:%22German%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:ger%20OR%20language:%22German%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
48. ^ "Internet Archive Search : (language:spa OR language:"Spanish")" ([https://archive.org/search.php?query=\(language:spa%20OR%20language:%22Spanish%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:spa%20OR%20language:%22Spanish%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
49. ^ "Internet Archive Search : (language:rus OR language:"Russian") AND mediatype:texts" ([https://archive.org/search.php?query=\(language:rus%20OR%20language:%22Russian%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:rus%20OR%20language:%22Russian%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
50. ^ "Internet Archive Search : (language:ara OR language:"Arabic")" ([https://archive.org/search.php?query=\(language:ara%20OR%20language:%22Arabic%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:ara%20OR%20language:%22Arabic%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
51. ^ "Internet Archive Search : (language:urd OR language:"Urdu") AND mediatype:texts" ([https://archive.org/search.php?query=\(language:urd%20OR%20language:%22Urdu%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:urd%20OR%20language:%22Urdu%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
52. ^ "Internet Archive Search : (language:Portuguese OR language:"por") AND mediatype:texts" ([https://archive.org/search.php?query=\(language:Portuguese%20OR%20language:%22por%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:Portuguese%20OR%20language:%22por%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
53. ^ "Internet Archive Search : (language:Chinese OR language:"chi") AND mediatype:texts" ([https://archive.org/search.php?query=\(language:Chinese%20OR%20language:%22chi%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:Chinese%20OR%20language:%22chi%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
54. ^ "Internet Archive Search : (language:Japanese OR language:"jpn") AND mediatype:texts" ([https://archive.org/search.php?query=\(language:Japanese%20OR%20language:%22jpn%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:Japanese%20OR%20language:%22jpn%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 21, 2014.
55. ^ "Internet Archive Search : (language:Dutch OR language:"dut") AND mediatype:texts" ([https://archive.org/search.php?query=\(language:Dutch%20OR%20language:%22dut%22\)%20AND%20mediatype:texts](https://archive.org/search.php?query=(language:Dutch%20OR%20language:%22dut%22)%20AND%20mediatype:texts)). Internet Archive. Retrieved January 23, 2014.
56. ^ "Internet Archive Search : mediatype:texts AND date:[1800-01-01 TO 1809-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1800-01-01%20TO%201809-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
57. ^ "Internet Archive Search : mediatype:texts AND date:[1810-01-01 TO 1819-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1810-01-01%20TO%201819-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
58. ^ "Internet Archive Search : mediatype:texts AND date:[1820-01-01 TO 1829-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1820-01-01%20TO%201829-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
59. ^ "Internet Archive Search : mediatype:texts AND date:[1830-01-01 TO 1839-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1830-01-01%20TO%201839-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
60. ^ "Internet Archive Search : mediatype:texts AND date:[1840-01-01 TO 1849-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1840-01-01%20TO%201849-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
61. ^ "Internet Archive Search : mediatype:texts AND date:[1850-01-01 TO 1859-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1850-01-01%20TO%201859-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
62. ^ "Internet Archive Search : mediatype:texts AND date:[1860-01-01 TO 1869-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1860-01-01%20TO%201869-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
63. ^ "Internet Archive Search : mediatype:texts AND date:[1870-01-01 TO 1879-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1870-01-01%20TO%201879-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
64. ^ "Internet Archive Search : mediatype:texts AND date:[1880-01-01 TO 1889-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1880-01-01%20TO%201889-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
65. ^ "Internet Archive Search : mediatype:texts AND date:[1890-01-01 TO 1899-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1890-01-01%20TO%201899-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
66. ^ "Internet Archive Search : mediatype:texts AND date:[1900-01-01 TO 1909-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1900-01-01%20TO%201909-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
67. ^ "Internet Archive Search : mediatype:texts AND date:[1910-01-01 TO 1919-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1910-01-01%20TO%201919-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
68. ^ "Internet Archive Search : mediatype:texts AND date:[1920-01-01 TO 1929-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1920-01-01%20TO%201929-12-31%5D>). Internet Archive. Retrieved January 24, 2014.

- query=mediatype%3A%28texts%29%20AND%20date%3A%5B1920-01-01%20TO%201929-12-31%5D). Internet Archive. Retrieved January 24, 2014.
69. ^ "Internet Archive Search : mediatype:texts AND date:[1930-01-01 TO 1939-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1930-01-01%20TO%201939-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
70. ^ "Internet Archive Search : mediatype:texts AND date:[1940-01-01 TO 1949-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1940-01-01%20TO%201949-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
71. ^ "Internet Archive Search : mediatype:texts AND date:[1950-01-01 TO 1959-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1950-01-01%20TO%201959-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
72. ^ "Internet Archive Search : mediatype:texts AND date:[1960-01-01 TO 1969-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1960-01-01%20TO%201969-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
73. ^ "Internet Archive Search : mediatype:texts AND date:[1970-01-01 TO 1979-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1970-01-01%20TO%201979-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
74. ^ "Internet Archive Search : mediatype:texts AND date:[1980-01-01 TO 1989-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1980-01-01%20TO%201989-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
75. ^ "Internet Archive Search : mediatype:texts AND date:[1990-01-01 TO 1999-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B1990-01-01%20TO%201999-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
76. ^ "Internet Archive Search : mediatype:texts AND date:[2000-01-01 TO 2009-12-31]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B2000-01-01%20TO%202009-12-31%5D>). Internet Archive. Retrieved January 24, 2014.
77. ^ "Internet Archive Search : mediatype:texts AND date:[2010-01-01 TO 2014-1-24]" (<https://archive.org/search.php?query=mediatype%3A%28texts%29%20AND%20date%3A%5B2010-01-01%20TO%202014-1-24%5D>). Internet Archive. Retrieved January 24, 2014.
78. ^ Gonsalves, Antone (December 20, 2006). "Internet Archive Claims Progress Against Google Library Initiative" (<http://www.informationweek.com/story/showArticle.jhtml?articleID=196701339>). InformationWeek.
79. ^ "The Open Library Makes Its Online Debut" (<http://web.archive.org/web/20070930184259/http://chronicle.com/wiredcampus/index.php?id=2235?atwc>). *The Wired Campus* (Chronicle of Higher Education). July 19, 2007. Archived from the original (<http://chronicle.com/wiredcampus/index.php?id=2235?atwc>) on September 30, 2007.
80. ^ "Search Inside" (<https://openlibrary.org/search/inside>) (feature). OpenLibrary.org.
81. ^ Internet Archive (June 25, 2011). "In-Library eBook Lending Program Expands to 1,000 Libraries" (<http://blog.archive.org/2011/06/25/in-library-ebook-lending-program-expands-to-1000-libraries/>). *Internet Archive Blogs*. June 25, 2011.
82. ^ "Internet Archive Search: Collection: Feature Films" (http://archive.org/search.php?query=collection%3Afeature_films&sort=%2Fadditional%2Fitem%2Fdownloads). Internet Archive. Retrieved March 6, 2013.
83. ^ "FedFlix" (<https://archive.org/details=FedFlix>). Internet Archive. Retrieved December 14, 2013.
84. ^ "September 11th Television Archive" (https://archive.org/details/sept_11_tv_archive). Internet Archive.]
85. ^ Aune, Sean (August 10, 2010). "The 40 best free legal movies you can download right now" (<http://tech.blorge.com/Structure:%20/2010/08/11/top-40-best-free-legal-movies-you-can-download-right-now/>). blorge.com.
86. ^ "Welcome to Machinima" (<https://archive.org/details/machinima>). Internet Archive.
87. ^ "TV NEWS : Search Captions. Borrow Broadcasts : TV Archive : Internet Archive" (<http://archive.org/details/tv/>). Internet Archive. Retrieved April 13, 2013.
88. ^ Fowler, Geoffrey A.; Hagey, Keach (September 18, 2012). "Let's Go to the Videotape: Nonprofit Offers News Clips" (<http://online.wsj.com/article/SB10000872396390443720204578002592487339454.html>). The Wall Street Journal Online.(subscription required)
89. ^ Kahle, Brewster (September 17, 2012). "Launch of TV News Search & Borrow with 350,000 Broadcasts" (<http://blog.archive.org/2012/09/17/launch-of-tv-news-search-borrow-with-350000-broadcasts>). *Internet Archive Blogs*.
90. ^ "Welcome to Audio Archive" (<https://archive.org/details/audio>). Internet Archive.
91. ^ Tirpack, Alex (June 3, 2009). "Warren Zevon live shows hit the web, possible film in the works" (<http://www.rollingstone.com/music/news/warren-zevon-live-shows-hit-the-web-possible-film-in-the-works-20090603>). *Rolling Stone*.
92. ^ "Welcome to Netlabels" (<https://archive.org/details/netlabels>). Internet Archive.
93. ^ Boswell, Wendy (October 21, 2006). "Download free music at the Internet Archive" (<http://liferhacker.com/208221/download-free-music-at-the-internet-archive>). *Liferhacker*. "The Internet Archive has a ginormous collection of free, downloadable music in their NetLabels category [...]"
94. ^ "NASA Images" (http://web.archive.org/web/*http://nasaimages.org/) (archive). Internet Archive. Retrieved April 13, 2013.
95. ^ "Internet Archive founder turns to new information storage device – the book" (<http://www.webcitation.org/6A683xLKE>). *The Guardian*. August 1, 2011. Archived from the original (<http://www.guardian.co.uk/books/2011/aug/01/internet-archive-books-brewster-kahle>) on August 21, 2012. "Brewster Kahle, the man behind a project to file every webpage, now wants to gather one copy of every published book"
96. ^ Scott, Jason. (April 13, 2013). "Change Computer History Forever: Well, Here We Are" (<http://ascii.textfiles.com/archives/3947>). Jason Scott's Weblog.
97. ^ Robertson, Adi (October 25, 2013). "The Internet Archive puts Atari games and obsolete software directly in your browser" (<http://www.theverge.com/2013/10/25/5028974/internet-archives-new-historic-software-collection>). The Verge.
98. ^ "Archive Page - What's the heck is going on at Internet Archive?" (<http://www.stevensaylor.com/ArchivePage.html>). Stevensaylor.com. January 8, 2013.
99. ^ "Unca Harlan's Art Deco Dining Pavilion: Another Can of Worms" (<http://harlanellison.com/heboard/archive/unca20130512.htm>). January 8, 2013 posting by Steve Perry. Harlanellison.com.
100. ^ "OMNI, Digital Rights, and Voodoo" (<http://harlanellison.com/heboard/archive/unca20130512.htm>). Posting of January 8, 2013. Harlanellison.com.
101. ^ Fleishman, Glenn (August 2, 2013). "The rebirth of Omni—and its vibe" (<http://boingboing.net/2013/08/02/omni.html>). BoingBoing.net.
102. ^ Mayshark, Jeff (December 1, 2005). "Wrath of Deadheads stalls a Web crackdown" (<http://www.nytimes.com/2005/12/01/technology/01iht-deadheads.html>). *The New York Times*.
103. ^ Lesh, Phil (November 30, 2005). "An Announcement from Phil Lesh" (http://web.archive.org/web/20070715072102/http://www.phillesh.net/philzonepages/friends_stuff/hotline-051130.html). *Hotline* (blog). PhilLesh.net. Archived from the original (http://www.phillesh.net/philzonepages/friends_stuff/hotline-051130.html) on July 15, 2007.
104. ^ Kahle, Brewster; Vernon, Matt (December 1, 2005). "Good News and an Apology: GD on the Internet Archive" (<http://www.archive.org/post/49553/good-news-and-an-apology-gd-on-the-internet-archive>). *Live Music Archive Forum*. Internet Archive. *Authors and date indicate the first posting in the forum thread*.
105. ^ Broache, Anne (May 7, 2008). "FBI rescinds secret order for Internet Archive records" (http://www.news.com/8301-10784_3-9938603-7.html). CNet.
106. ^ Nakashima, Ellen, "FBI Backs Off From Secret Order for Data After Lawsuit" (<http://www.washingtonpost.com/wp-dyn/content/article/2008/05/07/AR2008050703808.html>). Washington Post, May 8, 2008.
107. ^ Stalinsky, Steven (August 17, 2011). "Al-Qaeda, Jihadis Infest the San Francisco, California-Based 'Internet Archive' Library" (<http://www.memri.org/report/en/0/0/0/0/0/5576.htm>). The Middle East Media Research Institute.
108. ^ "Open Content Alliance" (<http://www.opencontentalliance.org/>). opencontentalliance.org. Retrieved April 13, 2013.
109. ^ Kahle, Brewster (January 17, 2012). "12 Hours Dark: Internet Archive vs. Censorship" (<http://blog.archive.org/2012/01/17/12-hours-dark-internet-archive-vs-censorship/>). *Internet Archive Blogs*.
110. ^ "Internet Archive Search : sponsor:'Yahoo!'" (<https://archive.org/search.php?query=sponsor%3A%22Yahoo%21%22>). Internet Archive. Retrieved March 1, 2014.
111. ^ "Internet Archive Search : sponsor:'msn'" (<https://archive.org/search.php?query=sponsor%3A%22msn%22>). Internet Archive. Retrieved March 1, 2014.
112. ^ "Internet Archive Search : sponsor:'Google'" (<https://archive.org/search.php?query=sponsor%3A%22Google%22>). Internet Archive. Retrieved March 1, 2014.
113. ^ "Internet Archive Search : sponsor:'Lyasis Members and Sloan Foundation'" (<https://archive.org/search.php?query=sponsor%3ALYRASIS%20Members%20and%20Sloan%20Foundation>). Internet Archive. Retrieved March 1, 2014.
114. ^ "Internet Archive Search : sponsor:'University of Toronto'" (<https://archive.org/search.php?query=sponsor%3A%22University%20of%20Toronto%22>). Internet Archive. Retrieved March 1, 2014.
115. ^ "Internet Archive Search : sponsor:'National Library of Scotland'" (<https://archive.org/search.php?query=sponsor%3A%22National%20Library%20of%20Scotland%22>). Internet Archive. Retrieved March 2, 2014.
116. ^ "Internet Archive Search : sponsor:'Natural History Museum Library, London'" (<https://archive.org/search.php?query=sponsor%3A%22Natural%20History%20Museum%20Library%20C%20London%22>). Internet Archive. Retrieved March 2, 2014.
117. ^ "Internet Archive Search : sponsor:'University of Alberta Libraries'" (<https://archive.org/search.php?query=sponsor%3A%22University%20of%20Alberta%20Libraries%22>). Internet Archive. Retrieved March 2, 2014.
118. ^ "Internet Archive Search : sponsor:'Research Library, The Getty Research Institute'" (<https://archive.org/search.php?query=sponsor%3A%22Research%20Library%20The%20Getty%20Research%20Institute%22>). Internet Archive. Retrieved March 2, 2014.
119. ^ "Internet Archive Search : sponsor:'Boston Library Consortium'" (<https://archive.org/search.php?query=sponsor%3A%22Boston%20Library%20Consortium%22>). Internet Archive. Retrieved March 2, 2014.

120. ^ "Internet Archive Search : sponsor:"The Library of Congress"" (<https://archive.org/search.php?query=sponsor%3A%22The%20Library%20of%20Congress%22>). Internet Archive. Retrieved March 2, 2014.
121. ^ "Internet Archive Search : sponsor:"Allen County Public Library"" (<https://archive.org/search.php?query=sponsor%3A%22Allen%20County%20Public%20Library%22>). Internet Archive. Retrieved March 2, 2014.
122. ^ "Internet Archive Search : sponsor:"Internet Archive" AND mediatype:texts" (<https://archive.org/search.php?query=sponsor%3A%22Internet%20Archive%22%20AND%20mediatype%3Atexts>). Internet Archive. Retrieved March 2, 2014.
123. ^ "Internet Archive Search : sponsor:"Harvard University" AND mediatype:texts" (<https://archive.org/search.php?query=sponsor%3A%22Harvard%20University%22%20AND%20mediatype%3Atexts>). Internet Archive. Retrieved March 2, 2014.
124. ^ "Internet Archive Search : sponsor:"China-America Digital Academic Library (CADAL)" AND mediatype:texts" (<https://archive.org/search.php?query=sponsor%3A%22China-America%20Digital%20Academic%20Library%20%28CADAL%29%22%20AND%20mediatype%3Atexts>). Internet Archive. Retrieved March 2, 2014.
125. ^ "Internet Archive Search : sponsor:"University of Illinois Urbana-Champaign" AND mediatype:texts" (<https://archive.org/search.php?query=sponsor%3A%22University%20of%20Illinois%20Urbana-Champaign%22%20AND%20mediatype%3Atexts>). Internet Archive. Retrieved March 2, 2014.
126. ^ "Internet Archive Search : sponsor:"University of North Carolina at Chapel Hill" AND mediatype:texts" (<https://archive.org/search.php?query=sponsor%3A%22University%20of%20North%20Carolina%20at%20Chapel%20Hill%22%20AND%20mediatype%3Atexts>). Internet Archive. Retrieved March 2, 2014.
127. ^ "Internet Archive Search : sponsor:"Biodiversity Heritage Library" AND mediatype:texts" (<https://archive.org/search.php?query=sponsor%3A%22Biodiversity%20Heritage%20Library%22%20AND%20mediatype%3Atexts>). Internet Archive. Retrieved March 3, 2014.

Further reading

- Kahle, Brewster, "Archiving the Internet" (https://web.archive.org/web/19971211123138/www.archive.org/sciam_article.html) at the Wayback Machine (archived December 11, 1997)
- Ringmar, Erik, "Liberate and Disseminate (<http://www.timeshighereducation.co.uk/story.asp?storycode=401386>)," *Times Higher Education Supplement*, April 10, 2008.

External links

- Official website (<https://archive.org>) in California, United States
 - Official blog (<http://blog.archive.org>)
- Internet Archive Mirror (<http://archive.bibalex.org/>) at the Bibliotheca Alexandrina, Egypt
- The *Netlabels* (<http://www.archive.org/details/netlabels>) collection at the **Internet Archive** ^[*more*]
- The *Internet Archive Press* (<http://www.archive.org/details/iapress>) collection at the **Internet Archive** ^[*more*]
- The *Internet Archive Presents* (<http://www.archive.org/details/internetarchivepresents>) collection at the **Internet Archive** ^[*more*]
- The Other Minds Archive (https://archive.org/details/other_minds) – New Music Resource from radiom.org (<http://radiom.org>)
- Web Archiving at archive.org (<http://iawebarchiving.wordpress.com/>), details of Internet Archive operations
- Pictures and descriptions of the Wayback Machine hardware in 2003 (prior to the Petabox), with cost information (<https://archive.org/web/hardware.php>)
- Current Petabox storage hardware (<https://archive.org/web/petabox.php>)
- Earliest known website (<https://web.archive.org/web/19971011064403/http://www.archive.org/index.html>) at the Wayback Machine (archived October 11, 1997)
- Early websites from 1996 (<http://classic-web.archive.org/collections/pioneers.html>)
- Brewster Kahle (November 6, 2013). "Scanning Center Fire — Please Help Rebuild" (<https://blog.archive.org/2013/11/06/scanning-center-fire-please-help-rebuild/>). *Internet Archive Blogs*.

Retrieved from "http://en.wikipedia.org/w/index.php?title=Internet_Archive&oldid=602973967"

Categories: 1996 establishments in California | Digital libraries | E-book suppliers | Film archives | Foundations based in the United States | History of the Internet | Internet properties established in 2001 | Non-profit organizations based in San Francisco, California | 501(c)(3) nonprofit organizations | Online archives | Organizations established in 1996 | Public libraries in California | Sound archives | Web archiving initiatives | Internet Archive projects | Mass digitization | Access to Knowledge movement

-
- This page was last modified on April 6, 2014 at 03:01.
 - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.